



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/488,028	01/20/2000	Eric Cohen-Solal	US 000015	3849

24737 7590 10/01/2003

PHILIPS INTELLECTUAL PROPERTY & STANDARDS  
P.O. BOX 3001  
BRIARCLIFF MANOR, NY 10510

EXAMINER

HAILU, TADESSE

ART UNIT	PAPER NUMBER
----------	--------------

2173

DATE MAILED: 10/01/2003

16

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

09/488,028

Applicant(s)

COHEN-SOLAL ET AL.

Examiner

Tadesse Hailu

Art Unit

2173

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☐ Responsive to communication(s) filed on 04 September 2003.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☐ Claim(s) 1,2,4 and 6-15 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☐ Claim(s) 1,2,4,6-10 and 12-15 is/are rejected.
- 7) ☐ Claim(s) 11 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on \_\_\_\_\_ is: a) ☐ approved b) ☐ disapproved by the Examiner.
- If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

**Priority under 35 U.S.C. §§ 119 and 120**

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
- a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

**Attachment(s)**

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) \_\_\_\_\_
- 4) ☐ Interview Summary (PTO-413) Paper No(s). \_\_\_\_\_
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_\_

**DETAILED ACTION**

1. This Office Action is in response to RCE Request entered 9/4/2003 for the patent application (09/488,028), filed on 1/20/2000.
2. The pending claims 1, 2, 4, 6-15 are examined.

***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

3. Claims 1, 2, 4, 6, 7, 12-15 are rejected under 35 U.S.C. 102(a) as being anticipated by Brøndsted, et al "The IntelliMedia WorkBench A Generic Environment For Multimodal Systems," (1998).

With regard to claim 1:

As per "a method of locating and displaying an image of a target," Brøndsted describes a method of locating and displaying an image of a target (see fig. 1);

As per "sensing a triggering event generated by a human operator," Brøndsted describes sensing spoken word (key word or command) as well as user's gesture via a microphone and camera respectively (see section 3);

As per "receiving additional external information that characterizes at least one machine-sensible feature of a target, said receiving step occurring substantially simultaneously with said sensing step;" since Brøndsted is a multimodal system, thus additional information about a target or location can be received through spoken word

(extracted key word) input as well as through gesture input (section 3). These inputs are executed simultaneously (section 2.1);

As per "aiming a camera in response to said sensing and said receiving step, wherein said sensing step includes sensing a gesture indicting a direction of said target." Brondsted describes a simultaneous speech and gesture input implemented on Workbench (see section 2.1). Brondsted further describes and illustrates (fig. 1) a camera directed toward the target; wherein the camera continuously captures the pointing hand over the workbench while the user/operator describes the location (section 2.1).

With regard to claim 2:

As per "... said sensing step includes sensing a gesture of a human operator indicating a target." Brondsted discloses Gesture recognizer (fig. 2) for sensing a gesture of a human operator indicating a target (see Brondsted, fig. 1).

With regard to claim 4:

As per "... said receiving step includes receiving speech from said human operator." Brondsted discloses Microphone (fig. 2) for receiving speech from said human operator (see Brondsted, section 2.1).

With regard to claim 6:

As per "... processing said speech for use with at least one machine sensor, said at least one machine sensor and said speech assisting in locating said target." Brondsted disclose Speech recognizer, Speech synthesizer, and Microscope (see Brondsted, fig. 2, and section 2.1).

With regard to claim 7:

As per "... said sensing step includes sensing a gesture indicting a direction from said human operator to said target." Brondsted discloses a gesture indicating a direction from said human operator to said target (see Brondsted, fig. 1).

With regard to claim 12:

As per "A method of locating and displaying an image of a target," Brondsted describes a method of locating and displaying an image of a target (see fig. 1);

As per "scanning an area within the range of at least one sensor;" Brondsted illustrates scanning a workbench (area) with a camera and laser pointer (fig. 1);

As per "identifying potential targets;" Brondsted describes identifying location (target) on the workbench shown with a campus information application (see section 2);

As per "storing information concerning machine sensible characteristics and locations of said possible targets;" Brondsted describes storing the multimodal campus information (section 2.1).

As per "sensing a triggering event, said triggering event generated by a human operator;" Brondsted describes sensing spoken word (key word or command as well as user's gesture via a microphone and camera respectively (see section 3);

As per "receiving additional external information that characterizes at least one feature of said target, said receiving step occurring substantially simultaneously with said sensing step;" since Brondsted is a multimodal system, thus additional information about a target or location can be received through spoken word (extracted key word)

input as well as through gesture input (section 3). These inputs are executed simultaneously (section 2.1); and

As per "aiming a camera in response to said sensing, storing and said receiving steps, wherein said sensing step includes sensing a gesture indicting a direction of said target." Brondsted describes a simultaneous speech and gesture input implemented on Workbench (see section 2.1). Brondsted further describes and illustrates (fig. 1) a camera directed toward the target; wherein the camera continuously captures the pointing hand over the workbench while the user/operator describes the location (section 2.1).

With regard to claim 13:

As per "A method of aiming a camera at a target," Brondsted illustrates aiming a camera and a laser pointer at a campus map location (target) (fig. 1).

As per "inputting an indication of a position of a target;" Brondsted illustrates and describes pointing toward a location of a target (fig. 1, see also section 3);

As per " inputting further information about a machine-sensible characteristic of said target;" Brondsted describes sensing spoken word (key word or command) as well as user's gesture via a microphone and camera respectively (see section 3);

As per "aiming a camera at said target in response to said indication and said further information to reduce an error in said aiming, wherein said inputting an indication step includes inputting a gesture indicting a direction of said target." Brondsted describes a camera and a laser pointer directed or focused toward a displayed workbench application. Brondsted further describes an automatic calibration procedure

involving both the camera and laser pointer. Using the camera the gesture recognizer is able to track 2D pointing gestures in real time. The camera continuously captures images. The hand gesture or the hand motion is analyzed in order to find the direction of the pointing hand (device) and its edge.

With regard to claim 14:

Claim 14, while not necessary identical in scope, contain limitations similar to independent claim 13 and therefore are rejected under the same rationale.

With regard to claim 15:

As per "...said step of orienting includes orienting a camera." Brøndsted, as illustrated in fig. 1 shows oriented camera view toward a workbench.

***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. Claims 8-10 are rejected under 35 U.S.C. 103(a) as being unpatentable over Tom Brøndsted, et al "The IntelliMedia WorkBench A Generic Environment For Multimodal Systems," (1998) in view of Indrajit Poddar, et al "Toward Natural Gesture/Speech HCI: A Case Study of Weather Narration," (1998).

With regard to claim 8:

As per "...said processing step includes processing said voice information through a look-up table corresponding said speech to search criteria for use with said at least one sensor." Brondsted does describe different module and for storing data, but Brondsted fails to describe "processing said voice information through a look-up table corresponding to said speech to search criteria for use with said at least one sensor." Poddar discloses a multimodal system, including speech (via Microphone) and gesture (hand) input (section 3). Poddar further discloses processing voice information through a look-up table (table1- table 4). Therefore it would have been obvious as the time the invention was made to replace Brondsted's voice information memory with Poddar's look-up table because it would be easier to structure the voice information and access the voice information as a table format.

With regard to claim 9:

As per "... said look-up table is modifiable." Brondsted in view of Poddar further describe replacing key words of the table, modifiable look-up table (Poddar, section 3).

With regard to claim 10:

As per "...said look-up table modifiable by receiving information through the on-line global compute network. ' since Brondsted system can be implemented in a distributed environment (sections 2.1- 2.2), the look-up table (voice data memory module) could be modified by information received from other remote devices.

***Allowable Subject Matter***



5. Claim 11 is objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The following is a statement of reasons for the indication of allowable subject matter: Although Brondsted and Poddar describes a modifiable look-up table (poddar, section 3) that includes replaces word or phrase input with another input and a corresponding search criteria (Poddar, section 3), "said added voice input and said corresponding search criteria established by comparing previous association of said added voice input with at least one machine sensible characteristic of at least one correctly identified target associated with said voice input, said machine sensible characteristic being a basis for determining said corresponding search criteria." not clearly described.

### ***Response to Arguments***

6. Applicant's arguments with respect to claims 1, 2, 4, and 6-15 have been considered but are moot in view of the new ground(s) of rejection.

### **Conclusion**

7. The prior art made of record on form PTO-892 and not relied upon is considered pertinent to applicant's disclosure. Applicant is required under 37 C.F.R § 1.111(c) to consider these reference fully when responding to this action. The documents cited therein teach Michael Johnston, et al describes a multimodal language processing architecture, which supports interfaces allowing simultaneous input from speech and gesture recognition. It is implemented in Quickset, a multimedia (pen/voice) system that

enables users to set up and control distributed interactive simulation (see Johnston, Abstract). Furthermore, in regard to the current independent claims, Johnston describes multimodal integration of spoken and gesture input accomplished simultaneously. The teaching of Johnston is implemented in QuickSet, a multimodal system that enables users to locate and display an image of a target, such as a specific location on a displayed map (page 1); Johnston also describes a sensing a command that is a speech describing spatial feature while simultaneously applying a gesture, such as marking the location on the map (page 1, page 4). To clearly understand the subject of interest which a user is gesturing or marking, Johnston describes a speech is simultaneously executed with gesture (page 4). However, Johnston does not describe the use of camera in associating to his teachings. Thus, Johnson fails to describe, "aiming a camera in response..." However, Bastian Leibe, et al describes the shortcomings, Bastian Leibe describes and illustrates (fig. 1) gesture input implemented on workbench. Bastian Leibe further describes a camera directed toward the target location the user is pointing to (fig. 1). Therefore, it would have been obvious at the time the present invention was made to incorporate the video camera of Bastian Leibe with Johnston's because senses and captures the operator/user's hand movement or gesture input remotely, which is an enhancement to Johnston's system.

8. Any inquiry concerning this communication or earlier communications from the Examiner should be directed to Tadesse Hailu, whose telephone number is (703) 306-2799. The Examiner can normally be reached on M-F from 10:00 - 8:30 ET. If attempts

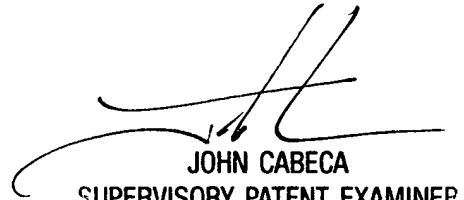
Art Unit: 2173

to reach the Examiner by telephone are unsuccessful, the Examiner's supervisor, John Cabeca, can be reached at (703) 308-3116 Art Unit 2173 CPK 2-4A51.

9. Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group receptionist whose telephone number is (703) 305-3900.

*Tadesse Hailu*

September 22, 2003



JOHN CABECA  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100